CENTRAL UNIVERSITY OF HARYANA
End Semester Examinations April 2022

Programme : M.Sc. Statistics
Semester : First
Course Title : Distribution Theory
Course Code : SBS ST 01 02 01 CC 4004

Session: 2021-2022
Max. Time : ***3 Hours***
Max. Marks : ***70***

1. Question no. 1 has seven parts and students need to answer any four. Each part carries three and half Marks.
2. Question no. 2 to 5 have three parts and student need to answer any two parts of each question. Each part carries seven marks.

Question No. 1. **(4X3.5=14)**

a) Obtain the skewness and kurtosis of a log-normal distribution.

b) Define central and non central t and χ^2 distributions.

c) Explain the bivariate normal distribution.

d) Define mixture distribution.

e) Arithmetic mean and standard deviation of a binomial distribution are respectively 4 and $\sqrt{8/3}$. Find the values of n and p.

e) 2% of the items made by a machine are defective. Find the probability that 3 or more items are defective in a sample of 100 items.

f) Explain compound and truncated distributions.

Question No. 2. **(2X7=14)**

a) For a trinomial distribution show that the correlation coefficient between X and Y is

$$\rho = -\sqrt{\left(\frac{p_1}{1-p_1}\right)\left(\frac{p_2}{1-p_2}\right)}.$$

b) Show that under certain conditions, negative binomial distribution tends to Poisson distribution.

c) Write down the probability mass function of binomial (n, p) distribution, if it is truncated at $x = 0$ and $x = n$.

Question No. 3. **(2X7=14)**

a) For a lognormal distribution, show that

$$E(X^r) = e^{r\mu + \frac{r^2\sigma^2}{2}}.$$

b)In an examination marks obtained by the student in Mathematics, Physics and Statistics are distributed normally about means 50, 52, and 18 with standard deviation is 15, 12,and 16 respectively. Find the probability of securing total marks of (a) 180 or above (b) 90 or below.

c) For a normal distribution $N(\mu, \sigma^2)$, obtain $E|X - \mu|$.

Question No. 4. **(2X7=14)**

a) Let X and Y are two iid binomial variate with parametera (n, p) and (m, p) respectively, then find the conditional distribution of $X \mid X+Y$.

b) Show that for t-distribution with k degrees of freedom

$$P[X \le x] = 1 - \frac{1}{2} I_{\frac{k}{k+x^2}}\left(\frac{k}{2}, \frac{1}{2}\right),$$

where

$$I_x(a,b) = \frac{1}{B(a,b)} \int_0^x y^{a-1}(1-y)^{b-1} dy$$

is the incomplete beta function.

c) Let X_i be independently distributed as $\chi^2_{n_i}$, $i = 1,2$. Obtain the probability density function of $Y = \frac{X_1}{X_2}$.

Question No. 5. **(2X7=14)**

a) Let X_1, X_2 denote a random sample of size 2 from a distribution $N(\mu, \sigma^2)$. Define $Y_1 = X_1 + X_2$ and $Y_2 = X_1 + 2X_2$. Show that the joint *pdf* of Y_1 and Y_2 is bivariate normal distribution with correlation coefficient $\frac{3}{\sqrt{10}}$.

b) Define $p-$variate normal distribution obtain its probability density function.

c) Show that if X_1 and X_2 are standard normal variates with correlation coefficient ρ between them, then the correlation coefficient between X_1^2 and X_2^2 is given by ρ^2.

CENTRAL UNIVERSITY OF HARYANA

Term End Examinations, April/May 2022

Programme: M.Sc (Statistics) **Session: 2021-22**

Semester: First **Max. Time: 3 Hours**

Course Title: Probability Theory **Max. Marks: 70**

Course Code: SBS ST01 102 C3104

Instructions:

1. Question no. 1 has seven sub parts and students need to answer any four. Each sub part carries three and half Marks.

2. Question no. 2 to 5 have three sub parts and students need to answer any two sub parts of each question. Each sub part carries seven marks.

Question No. 1. (4X3.5=14)

a) Define the following terms: (i) σ -field and (ii) Borel σ -field.

b) If X and Y are independent random variables, then show that $E(XY)=E(X)E(Y)$.

c) Prove that if A and B are independent events then $\overline{A}$ and $\overline{B}$ are also independent events.

d) Describe explicitly the sample spaces for each of the following experiments:

(i) The tossing of four coins.

(ii) The throwing of three dice.

e) If $h(x)$ be a non-negative borel-measurable function of an random variable X, if $E(h(x))$ exist then show that for $\varepsilon>0$, $P[h(x)\geq\varepsilon]\leq\frac{E[h(x)]}{\varepsilon}$.

f) Define following law of large number

(i) WLLN

(ii) SLLN

g) Define the quantile of order "p". Also obtain the quantile of order p (0< p <1) for the pdf

$$f(x)=\begin{cases}\frac{1}{x^2}, & x\geq 1\\ 0 & otherwise.\end{cases}$$

Question No. 2. (2X7=14)

a) State and prove the Boole's inequality.

b) If A, B and C are mutually independent events then $A\cup B$ and C are also independent events.

c) Sate and prove Bayes theorem.

Question No. 3. (2X7=14)

a) The Joint pdf of two random variables X and Y is given by

$$f(x,y)=\frac{9(1+x+y)}{2(1+x)^4(1+y)^4}; \qquad 0\leq x<\infty,\ 0<y<\infty.$$

Find the marginal distributions of X and Y and the conditional distribution of Y for $X=x$.

b) Let X be non-negative random variable with distribution function F then show that $E(X)=\int_0^\infty[1-F(x)]dx$.

c) If X is a random variable with mean 0 and standard deviation σ. Then Show that

$$F(t) \leq \frac{\sigma^2}{\sigma^2 + t^2};\ t < 0,$$

$$F(t) \geq \frac{\sigma^2}{\sigma^2 + t^2};\ t > 0.$$

where $F(t) = P(X \leq t)$.

Question No. 4. (2X7=14)

a) Define moment generating function and characteristic function and present their properties.
b) State and prove the Inversion theorem.
c) Find the density function $f(x)$ corresponding to characteristic function defined as follows:

$$\phi(t) = \begin{cases} 1 - |t|, & |t| \leq 1 \\ 0, & |t| > 1 \end{cases} .$$

Question No. 5. (2X7=14)

a) Define the statement of central limit theorem(CLT) and present the Laypounov's condition for sequence of independent random variable $\{X_i\}$ to hold CLT. Examine whether CLT hold for following:

(i) $P(X_k = \pm\sqrt{k}) = \frac{1}{2}$

(ii) $P(X_k = \pm k) = \frac{1}{2} k^{-\left(\frac{1}{2}\right)}$, $P(X_k = 0) = 1 - k^{-\left(\frac{1}{2}\right)}$.

b) Prove that if a sequence of a random variable $\{X_n\}$ converges in the rth mean then it also converges in probability, Symbolically, $X_n \xrightarrow{r^{th}} X \Rightarrow X_n \xrightarrow{P} X$.

c) State and prove the Chebyshev's inequality.

CENTRAL UNIVERSITY OF HARYANA

Term End Examination April 2022

Programme	**: M.Sc. Statistics**	**Session: 2021-22**
Semester	**: First**	*Max. Time:* ***3 Hours***
Course Title	**: Introductory Statistics**	*Max. Marks:* ***50***
Course Code	**: SBS ST 01 101 GE 3104**	

Instructions:

1. Question no. 1 has five parts and students need to answer any four. Each part carries two and half Marks.
2. Question no. 2 to 5 have three parts and student need to answer any two parts of each question. Each part carries five marks.

Note: Scientific Non programmable Calculator is allowed

Question No. 1. **(4x2.5=10)**

a. Discuss the meaning of level of significance with suitable example.

b. Point out the difference between the simple arithmetic mean and weighted arithmetic mean.

c. What are the characteristics of good measure of dispersion?

d. What is a Binomial distribution? Give a real life example where such a distribution is appropriate.

e. Define any two types of scale of measurement.

Question No. 2 **(2x5=10)**

a. Define a random variable. Distinguish between discrete and continues random variable. Define distribution function of a random variable and state its important properties.

b. The table below gives the diastolic blood pressure of 250 men. The readings were made to tne nearest millimetre and the central value of each group is given below. Calculate the median from the data.

Blood Pressure (mm)	60	65	70	75	80	85	90	95
No. of Men	4	5	31	39	114	30	25	2

c. Define primary and secondary data. What are the main sources of secondary data?

Question No. 3 **(2x5=10)**

a. Define Probability Mass function. Is the following:

$$\begin{cases} 2x, & 0 < x \le 1 \\ 4-2x, & 1 < x \le 2 \\ 0, & \text{elsewhere} \end{cases}$$

a probability density function?

b. In 10 independent throws of a defective dice, the probability that an even number will appear 5 times is twice the probability that an even number will appear 4 times. Find the probability that an even number will not appear at all in 10 independent throws of the dice.

c. Define Normal distribution. What are the main characteristics of a Normal distribution?

Question No. 4 **(2x5=10)**

a. Explain the meaning and uses of analysis of variance. How is a one way analysis of variance table setup? What are the assumptions of this test?

b. In a sample of 8 observations, the sum of squared deviation of items from the mean was 84.4. In another sample of 10 observations, the value was found to be 102.6. Test whether the difference is significant at 5% level.

You are given that at 5% level, critical value of F for v_1= 7 and v_2= 9 degrees of freedom is 3.29 and for v_1= 8 and v_2= 10 degrees of freedom, its value is 3.07.

c. The life time of electric bulbs for a random sample of 10 from a large consignment gave the following data:

Item	1	2	3	4	5	6	7	8	9	10
Life in '000 hours	4.2	4.6	3.9	4.1	5.2	3.8	3.9	4.3	4.4	5.6

Can we accept the hypothesis that the average life time of bulb is 4,000 hours

The value of t at 5 % level of significance for v= 10 is 2.22

The value of t at 5 % level of significance for v= 9 is 2.26

Question No. 5 **(2x5=10)**

a. Define correlation and explain its importance in statistical analysis.

b. Explain the concept of regression and point out its significance. Discuss some applications of regression analysis.

c. In the following table are recorded data showing the test scores made by 10 salesmen on an intelligence test and their weekly sales:

Salesmen	1	2	3	4	5	6	7	8	9	10
Test Scores	50	70	50	60	80	50	90	50	60	60
Sales ('000 Rs)	25	60	45	50	45	20	55	30	45	30

Calculate the rank correlation coefficient between intelligence and efficiency in salesmanship.

CENTRAL UNIVERSITY OF HARYANA

End Semester Examinations April 2022

Programme:	**M.Sc. Statistics**	**Session: 2021-22**
Semester:	**First**	**Max. Time: 3 Hours**
Course Title:	**Sampling Techniques**	**Max. Marks: 70**
Course Code:	**SBS ST 01 104 C 3104**	

Instructions:

1. Question no. 1 has seven parts and students need to answer any four. Each part carries three and half Marks.

2. Question no. 2 to 5 have three parts and student need to answer any two parts of each question. Each part carries seven marks.

Q 1. (4X3.5=14)

a) Show that in SRSWOR, the probability of selecting a specified unit of the population at any given draw is equal to the probability of its being selected at the first draw.

b) Mention the advantages and disadvantages of sample survey.

c) What is the purpose of stratification? Mention the principles of stratification.

d) Differentiate between linear and circular systematic sampling. Mention the situations where systematic sampling is used.

e) Differentiate between ratio and regression methods of estimation. In SRS show that the bias of the ratio estimator is given by $B(\hat{R}) = -\frac{Cov(\hat{R},\bar{x})}{\bar{X}}$.

f) Differentiate between two phase and two stage sampling with the help of example.

g) Explain Lahiri's method with example to select a sample with varying probabilities.

Q 2. (2X7=14)

a) Explain the principal steps involved in the planning and execution of sample survey.

b) In SRSWOR, show that the variance of the sample mean is given by $Var(\bar{y}_n) = \left(\frac{1}{n} - \frac{1}{N}\right) S^2$.

c) If $(X_i, Y_i); i = 1, 2, \ldots, N$ are the pairs of variates defined for every unit $i = 1, 2, \ldots, N$ of the population and $(\bar{x}_n, \bar{y}_n)$ are the corresponding means of the simple random sample $(x_i, y_i); i = 1, 2, \ldots, n$ of size n taken without replacement, then prove that $Cov(\bar{x}_n, \bar{y}_n) = \left(\frac{1}{n} - \frac{1}{N}\right)\frac{1}{N-1}\sum_{i=1}^{n}(X_i - \bar{X}_N)(Y_i - \bar{Y}_N)$, where $\bar{X}_N$ and $\bar{Y}_N$ are the corresponding population means.

Q3. (2X7=14)

a) In stratified random sampling with given cost function as $C = a + \sum_{i=1}^{k} C_i n_i$, derive the condition for variance $\bar{y}_{st}$ to be minimum.

b) If the population correction factor is ignored, then prove that

$$Var(\bar{y}_{st})_{opt} \leq Var(\bar{y}_{st})_{prop} \leq Var(\bar{y}\ \ )_{srs}.$$

c) In usual notations, show that in systematic sampling $Var(\bar{y}_{sys}) = \frac{nk-1}{nk}\frac{S^2}{n}\{1 + (n-1)\rho\}$, where ρ is the intraclass correlation coefficient between the units of the same systematic sample. Also, derive the condition under which systematic sampling is more efficient as compared to SRSWOR.

Q 4. (2X7=14)

a) Show that in SRSWOR, the variance of ratio estimator $\hat{R}$ is given by $Var(\hat{R}) = \frac{N-n}{nN}R^2[C_y^2 + C_x^2 + 2\rho C_x C_y]$. Also show that ratio estimate will be better than SRSWOR if $\rho > \frac{C_x}{2C_y}$.

b) Differentiate between separate and combined regression estimators. If $\bar{y}_{lc}$ denotes combined regression estimator then obtain variance of $\bar{y}_{lc}$, if sampling is independent in different strata and sample size is large enough in each stratum.

c) Show that the relative efficiency of cluster sampling with SRSWOR is given by $[1 + (M-1)\rho]^{-1}$, where ρ is the intra-cluster correlation coefficient. Also derive the condition when cluster sampling is more efficient than SRS.

Q 5. (2X7=14)

a) Define Horvitz Thompson estimator and obtain its variance. Also mention its drawbacks.

b) Let there be n fsu's and m ssu's are chosen from fsu's by SRSWOR, then show that $\bar{y}$, the sample mean per element is an unbiased estimate of population mean. Also obtain its variance.

c) Define double sampling for ratio estimator. Show that the ratio estimator of $\bar{y}$ under double sampling is biased and obtain its variance.